[*Translating and the Computer 9*. Proceedings of a conference… 12-13 November 1987, ed. Catriona Picken (London: Aslib, 1988)]

Machine translation: a threat or a promise?

Magnar Brekke

English Department, University of Bergen, Norway

Roald Skarsten

Arts Faculty Computing Section, University of Bergen, Norway

INTRODUCTION

Any elementary textbook on linguistics will tell us that, strictly speaking, there is no such thing as a synonym. And, by the same token, we may safely assume that there is no such thing as translation. It is by definition impossible, and when it is done, it is the task of the translator to minimise the damage.

There are undoubtedly greater challenges in translation than going between two relatively closely related languages like English and Norwegian, and yet, as many of us have found, the more you mingle with close relatives, the more false friends you tend to acquire or discover.

That risk you may be willing to take – except when your own life depends on the clarity and correctness of information communicated at all levels of an enormous industrial organisation constructing and operating a gigantic oil platform in the North Sea.

To many people at this Conference the name Bergen means – well, perhaps not as many of its inhabitants would like to think, Edvard Grieg and his music, or, indeed, the 1986 Eurovision Song Contest – but *oil terminology* and particularly a certain termbank for such a substance, developed at the University of Bergen over a period of several years, at the initiative and with substantial financial backing of Statoil. As part of this activity, the English department in conjunction with the Liberal Arts computing section became engaged in translating an entire library of quality assurance documentation between English and Norwegian, on the basis of terms provided by the Norwegian termbank and money provided by our three Norwegian oil companies.

When that project, named PETTRA, ground to a halt after about eighteen months (coinciding with the nose dive in the crude oil price in 1986 we had

gained a tremendous amount of experience – not only in tackling the outrageous demands from big customers with all deadlines yesterday, or in defending our technologically naive but linguistically impeccable principles against the onslaught of the engineers and vice versa, but above all, we gained experience in surviving the drudgery and tediousness of handling, with a modest degree of electronic word processing and look-up facilities, the stereotype text of technical specifications for the petroleum industry.

Very early in the game we became convinced that there had to be a more efficient way of doing this, and started asking around: whatever happened to machine translation after its near-demise in the mid-1960s?

The answer came back: MT is indeed alive and well and on its way to a veritable renaissance – but sorry, folks, Norwegian is not on the list! To cut a long story short, we managed in about a year to establish Project ENTRA, based on cooperation between

—Weidner Communications Corporation, Chicago, which allowed us to use its English-German MacroCAT as a starting point for an English-Norwegian version; and
—Digital Equipment Corporation of Norway, which provided free use of its brand new MicroVAX II for twelve months; and
—The University of Bergen, whose Directors would like to encourage joint projects where linguistic research and practical language competence could interface with the needs of the commercial/industrial community.

After a brief training session in Chicago 'Team ENTRA' was ready for action: during a 12-month period, on a part-time basis, the four of us (in addition to the two authors, the project team consists of Margaret Stenersen and Jon Erik Hagen) would pool our respective competences in computer programming, computational linguistics, English and Norwegian linguistics, as well as in practical translation, to produce the first operational system for automatic translation from English to a Nordic language. We have recently reached our goal and would like to share with you some of the insights we think we've gained, some of the problems we have tried to solve, and some of the perspectives that this work has given us on the current and future prospects of MT.

MODIFICATIONS OF THE ENGLISH-GERMAN MACROCAT

The following broad outline gives some indication of the types of tasks and operations we had to engage in:

1. English-German dictionary converted to English-Norwegian dictionary
2. Norwegian inflection rules systematised and programmed
3. Norwegian verb string equivalents worked out and entered in table
4. uniquely German rules deleted

5. Norwegian rules added for: (a) reordering; (b) insertion; (c) deletion; and (d) modifications:

(a) Reordering. In general terms, this involved changing sentence structure from a subject-verb language (English) to a verb-subject language (Norwegian):

Today I met him → Idag møtte jeg ham

Likewise there were systematic differences in adverb positions between the two languages:

I always go to Florida in the winter → Jeg reiser alltid til Florida om vinteren

Constructions with of tend to mask a variety of relations between the two nouns involved; for pragmatic reasons we have decided to turn them into a genitival construction:

The expansion of the refinery → raffineriets ekspansjon

Prepositions are, as everyone knows, the sort of thing you should not end an English sentence with . . . however, 'preposition stranding', as the phenomenon is called, happens to be obligatory in Norwegian, hence:

The chair on which you are sitting → Stolen som du sitter på

(b) Insertion. This proved necessary in a number of cases where English has various constructions available but Norwegian has only one or a different type altogether. Postmodifying *ing*-clauses thus had to be turned into an explicit relative clause (which is available even for English), while postmodifying adjectives, which do not occur in Norwegian, had to be preposed:

Dogs biting children → Hunder som biter barn
People starving → Sultende mennesker

The Norwegian infinitive of purpose must contain an explicit *for* in front of the infinitival marker a:

I bought the gun to shoot him → Jeg kjøpte pistolen for å skyte ham

As a final example of cases where Norwegian needs to be more explicit than English and therefore requires the insertion of an element, is the group of so-called transitive/reflexive verbs:

He shaves every morning → Han barberer seg hver morgen

(c) Deletions. The structurally most important operation here consists in getting rid of the preposed definite article when it occurs immediately before the noun; if an adjective intervenes, the article stays:

The big car → Den store bilen
the car → bilen

(d) Modifications. Some of these appear quite trivial, but require rather sophisticated solutions. Prepositions are an obvious case of strongly context-dependent selection:

By next Friday → INNEN neste fredag
By the river → VED elven

Other modifications involve English-Norwegian differences in syntactic patterns and relationships; for instance, Norwegian lacks a construction corresponding to 'raising', which consequently has to be 'unraised':

I want him to go → Jeg vil at han skal gå
'I want that he shall go'

Norwegian predicative adjectives need to show gender/number agreement with the subject:

The car is big → Bilen er stor
The house is big → Huset er stort
The cars are big → Bilene er store

A final example of modification concerns an important contrast in the way the two languages form complex noun phrases (NPs); like German, Norwegian concatenates:

pump room ventilation → pumperomventilasjon

As we will show below, concatenation raises fundamental questions as to the borderline between syntax and word formation and cannot be fully implemented before the linguistic rules have been worked out.

As a result of our efforts to implement these various syntactic operations we now have an English-Norwegian MacroCAT up and running, doing most of the things we have programmed it to do (as well as a few that we have not) and performing reasonably well in the various test runs and demonstrations that we perform for potential users. The system is in the process of being prepared for QA-testing at the Chicago headquarters and should be available on the open market early in 1988.

RESULTS OF INFORMAL ERROR ANALYSIS OF SAMPLES FROM THREE DIFFERENT TEXTS

At this point we would like to present the results of a pilot study of how good our MT system is at the moment, a status report from an ongoing error analysis which, when completed, will result in a graduate thesis. All the standard caveats regarding error analysis apply here, the more so because all figures are preliminary and tentative. Nevertheless, we dare to present them here because we believe they are better suited than a glorified and biased description to give

you some feel for just how far we have come at the moment and how much remains for us to do to improve the performance of this first version.

As a benchmark corpus for our pilot test we selected three documents with a total of about 7,500 words; one representing the petroleum sublanguage that was the basis of our dictionary work, another representing the (currently) unfamiliar sublanguage of a computer manual, and a third document representing the related but so far untested sublanguage of maritime fire protection regulations. Averaging our findings for these three documents, here is our main conclusion:

15 per cent of all raw-output sentences were flawless

Instead of speculating on the value of whatever criteria could be objectively applied in arriving at this conclusion, we offer below a brief analysis of the distribution of errors in the remaining 85 per cent of the sentences in the tested material. It must be repeated that error analysis is a notoriously slippery undertaking and that the entities dealt with have a very relative existence. Many errors are thus interdependent and could be hard to trace to their real source. No attempt has been made to grade the errors according to degree of severity, and accordingly, a given label or figure will cover a multitude of sins.

Within the 85 per cent of the sentences containing various errors we found the following distribution (shown in percentages of total number of errors):

6 per cent system errors:

Some of these relate to the behaviour of our extra Norwegian characters æ/ø/å inside idioms, a problem which is being eliminated.

43 per cent unsuitable analysis:

From the point of view of a post-editor, about two-thirds of these errors would be easy to correct, one-third relatively hard. Most of these derived from originals with very complex and often coordinated verb phrases, ambiguities or plain bad English. In the last two cases it would seem unreasonable to expect the machine to unravel what no human reader would be able to understand; perhaps one could argue here for persuading the technical author to use one of the existing text critiquing systems prior to sending off the original. Among minor problems we listed such errors as ‘wrong homograph’ and confusion of present tense forms and imperatives.

24 per cent dictionary errors:

After discounting the relatively few human errors detected in incorrect flagging of some feature there remains a sizable group under the category ‘missing word/term/idiom’. In the majority of cases the error derived from a lexical choice being made from the wrong sublanguage; as further evidence of this, a

dramatic increase was seen in this category when we tested another text from a totally new domain.

20 per cent programming errors:

These were of various kinds and derived chiefly from lack of time to supply the required code or from inability to determine, for the time being, choice of strategy.

7 per cent unsolved linguistic problems:

This percentage seems fairly modest in comparison with those above and could be an indication, perhaps a surprising one, that this is not where current MT is most vulnerable. Since the transfer difficulties arising in the interface between the two languages have been little studied in the amount of detail needed for MT, we will now present some of the more intriguing puzzles that we have not quite been able to solve so far:

1. **Idiomatic versus literal meaning:** In general, idioms can be handled quite adequately by being entered in the MacroCAT lexicon; however, in cases where each lexical element carries its literal meaning, there are no obvious ways in which the system can avoid them, any more than there are specific clues telling the human reader that the idiomatic reading does not apply. The following standard example serves to illustrate the point:

The pump operator kicked the bucket overboard

2. **Idiomatic choice of preposition:** This has already been alluded to under ‘modifications’ above. Prepositions are well-known stumbling blocks for any language learner and present major problems for any MT system making use of lexical substitutions. Another illustration of the magnitude of the problem:

Eng. *IN*	→	Norw.	*I,*
			INNEN,
			PÅ,
			OM,
			VED (Å),

3. **Multiple concatenations:** This phenomenon, which lies at the heart of Germanic NP-constructions, is clearly restricted with respect to the number of elements that can be concatenated, but the question of where Norwegian stops concatenating and begins using prepositional phrases or other syntactic means has not really been answered. The following example seems to indicate that four elements is a bit much:

Exhaust valve design criteria
→ ?eksosventilkonstruksjonskriterier
→ kriterier for konstruksjon av eksosventil

4. Imperatives: This problem also derives from the structure of complex NPs in English. These occur very frequently as headings and in lists in technical documentation. Since the parser gives priority to clauses over phrases, it will often interpret complex NPs as subjectless clauses, i.e. imperatives. In the language of instruction manuals this is quite appropriate, but in descriptions/specifications imperatives are out of place:

Control functions
→ *kontroller funksjoner
→ kontrollfunksjoner

5. Nonfinite verb forms: These are the chameleons of the language, particularly of English. They tend to mask a number of distinctions which a target language like Norwegian needs to make explicit. 'Purpose infinitives' have already been mentioned, but the greatest variety of problems turn up with the participial forms. Let us look at a few selected examples where, faced with alternative translations of equal linguistic plausibility, we have to make a choice in favour of 'maximum generality' in relation to the general language type we are dealing with:

ING-forms:
the cementing system
→ sementeringssystemet
→ det sementerende systemet

recirculating the fluid
→ resirkulering av væsken
→ å resirkulere væsken

the fire extinguishing arrangements
→ brannslokkingsopplegget
→ *brannen som slokker arrangementer

ED-forms:
bolts mounted on the outside
→ *bolter monterte (PRET.) på utsiden
→ bolter som er montert på utsiden

This final example affords us an opportunity to point out that sometimes an incorrect or inadequate analysis may yield an apparently correct result: mistaking the perfect form for the preterite becomes visible only when these are distinct (as with most strong verbs) in the target language.

THREAT OR PROMISE?

Before this experienced, knowledgeable and thoroughly professional audience we see no need to spend much time addressing the question whether or not MT is a threat. A human translator producing text that could be mistaken for MT raw output should be looking for other employment, anyway, but as is the case with any newly acquired tool, we need to be very explicit about the premises for its use. It should be eminently clear from our presentation that the current generation of MT software requires a linguistically informed, perhaps even sophisticated, environment – it is not a gimmick designed for secretaries and office supervisors, but a rather fragile productivity instrument which requires refinement and fine-tuning by a professional language person in order to yield the expected results.

Let me now, by way of conclusion, remind you of some of the reasons why we, based on our experience, think that MT in its proper setting represents a considerable promise for the information society of the future and for the information processing profession in particular.

Time

Since most text is already available in machine readable form, submitting it to a preliminary MT would in itself be a routine operation in most office environments. However, this could quickly become useless or counter-productive unless preceded by a look-up in a subject/domain relevant dictionary, which of course does not spring into being out of nowhere – it must have been carefully structured and accumulated by a skilled person with access to real textual and lexical data. The same (type of) person should also be responsible for entering unfound words and multi-word terms in the dictionary prior to batch translation, and for post-editing the machine output.

We will not here enter into the discussion about which criteria would be suitable for evaluating MT or for assessing the time saved or wasted by this or that system. We only observe that operational systems appear to save a non-trivial proportion of translator time,[1] and that this saving may play a crucial role in reducing the time-to-market for any new product out there in the commercial/industrial world.

Volume

This is of course only another aspect of the time factor. The growth in documentation seems to be near-exponential. A recent newspaper report stated that the number of documents required to operate a modern oil rig has increased tenfold in just a few years time and that the sheer weight of the paper was a factor to be reckoned with in calculating the stability of the platform! CD-ROM technology may of course alleviate that problem, but the desired or required availability of the same document in more than one language remains a

substantial challenge to the capacity of the translation profession. The paper mountain is simply insurmountable without the computer!

Cost

Since this is only a function of the previous two, very little needs to be said about it. Increasingly aggressive competition and price-cutting will force companies to maximise their efficiency of operation, and any instrument likely to contribute to cost reduction will be embraced. This could represent a real temptation for the marketing division of any MT vendor – but I am convinced that any sale of MT software which fails to point out the essential role of the translator will be deeply regretted. The initial investment in terms of both money, time, and effort is large, but it cannot be bypassed. Only after a considerable period of textual profiling and adaptation will the investment yield a substantial return – but it will!

Quality

So – are we going to ignore the question of quality? No, of course not – but as everyone knows, this is a highly relative and multifaceted concept. Despite the unquestionable shortcomings and the inevitable regressions experienced with fourth generation MT, the quality of raw output is improving. And the quality of the post-edited output – well, that's up to you, ladies and gentlemen.

REFERENCE

[1] See W.W. Cressey: Project for Integrated Development of English-Spanish Machine Translation. Review article in *Computers and Translation* (1987) 2 (1).